Name ______________________________ Date______________ Class____________

WORKSHEET

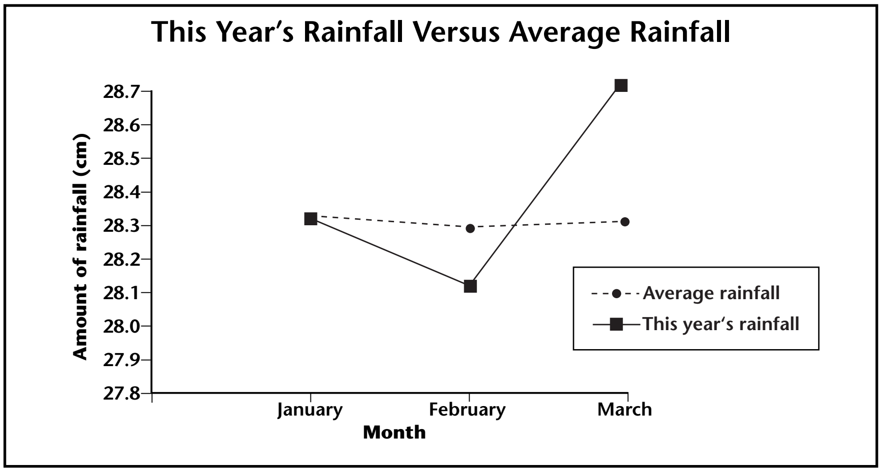
28 COMMUNICATING SKILLS

Recognizing Bias in Graphs

Graphs can be used to display your data at a glance. However, graphs can distort your results if you are not careful. The picture that results may not be **objective**, or without bias or distortion. Look at the first graph.

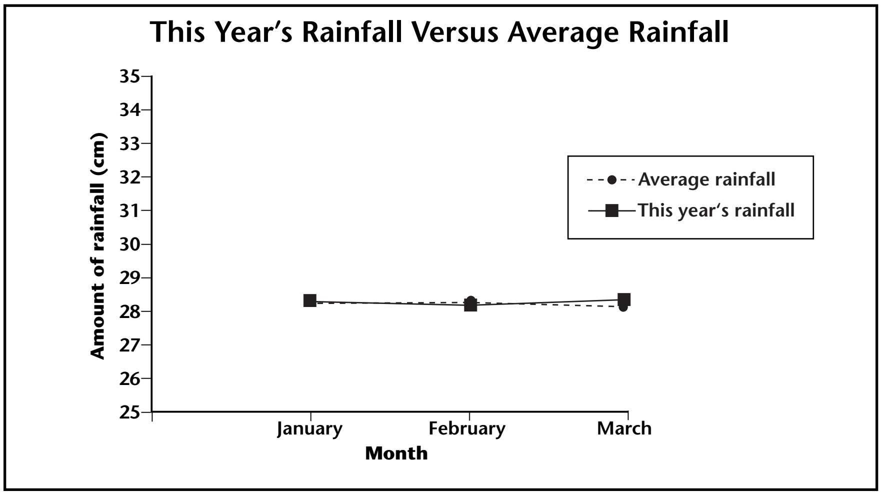
How Much Rain Really Fell?

In the graph below, it appears as though March had drastically more rainfall compared with an average month. But did that really happen?



Wait! March's rainfall was only 0.4 cm above average. On the graph, that looks like a large increase. On the ground, a 0.4 cm increase is not that much. This graph is *biased* because it exaggerates the difference between the two lines. Because the interval between 27.8 cm to 28.7 cm on the *y*-axis is so small, the difference in rainfall seems very large and noticeable.

If you increase the interval between numbers on the *y*-axis, the scale becomes larger. That makes the difference between the two lines smaller, as shown below.



Refer to the graphs on the previous page to answer the following questions:

1. What is the range of values on the *y*-axis in the second graph?

____

2. How does the difference between the two lines in the second graph compare with the difference between the two lines in the first graph? Which graph is a more accurate picture of the data? Explain.

____

____

____

____

A Matter of Scale

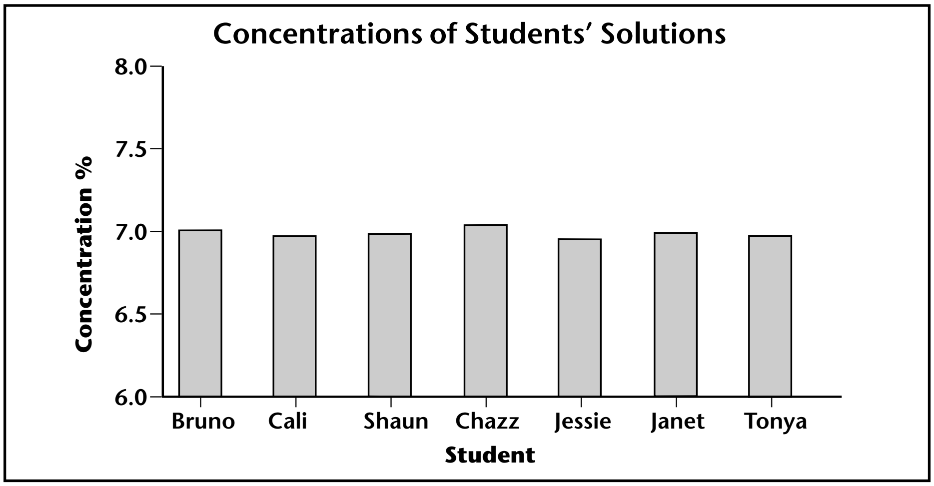
Here is another example of how the choice of the scale can alter a graph.

In an experiment, seven students tried to mix a solution of salt water so that its concentration would be exactly 7.00%. When the teacher tested the concentration of their solutions, he got the following results:

Concentrations of Students' Solutions

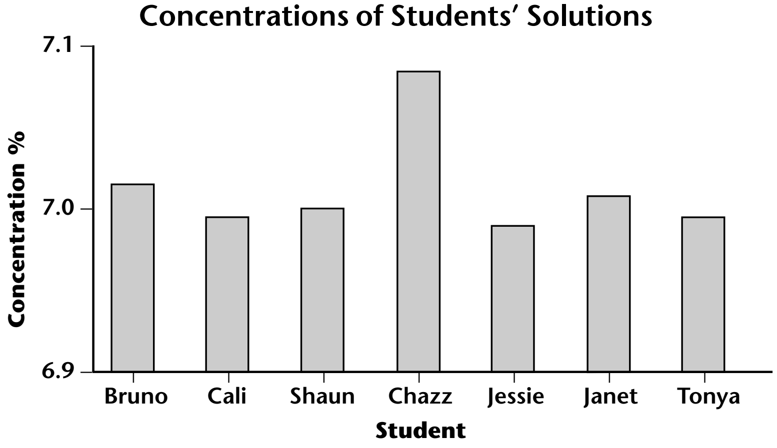
Name	Bruno	Cali	Shaun	Chazz	Jessie	Janet	Tonya
Concentration	7.02%	6.99%	7.00%	7.08%	6.97%	7.01%	6.99%

The teacher created the following graph to show the students' results:



Does this graph give you a clear picture of how the concentrations varied? Not really. The bars look so much alike that it's hard to tell the differences between them.

Suppose the teacher decreased the scale of the *y*-axis. The graph would then look like the one below. The variation in the students' results looks much greater, even though it hasn't changed. This graph makes it easier to see the small differences.



Graphs with an Attitude

The data in the chart below were recorded by a student measuring the thickness of four rock layers.

Thickness of Rock Layers

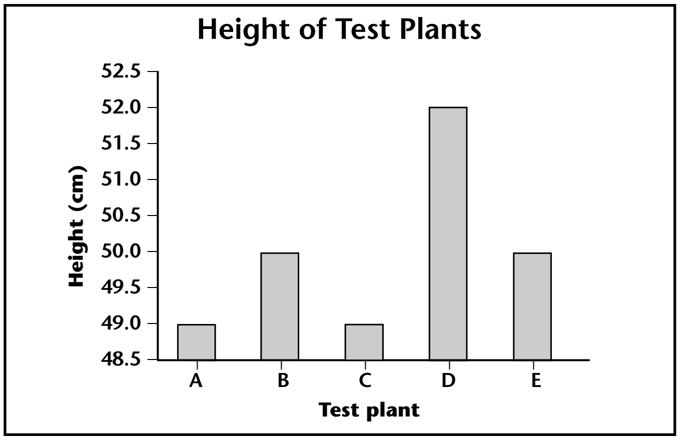
Layer	A	B	C	D
Thickness	11.2	10.8	13.5	11.1

Using the above data, create two graphs in the space below. First show how similar the measurements are. (Hint: Make the scale of the *y*-axis larger. This makes the difference between the measurements seem smaller.) In your second graph, emphasize the fact that layer C was slightly thicker than the other layers.

Your Graphs:

Identifying Bias on Your Own

Graph 1



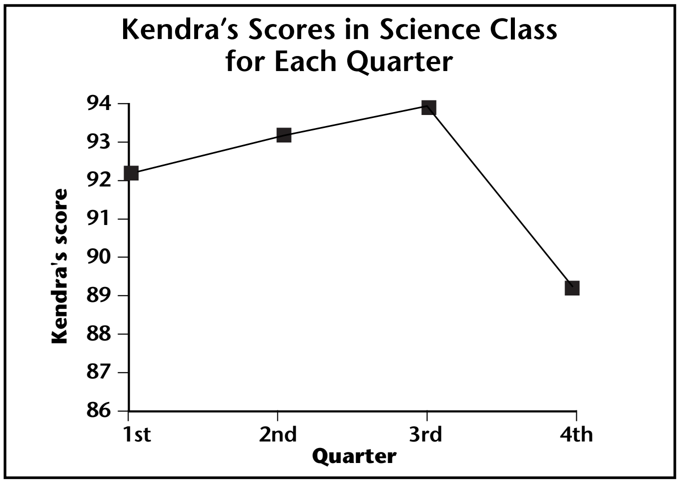
1. This graph shows that test plant D grew much taller than the other plants. How is this information misleading?

____

____

____

Graph 2



2. This graph shows that Kendra received a much lower grade in science class during the fourth quarter. Do you think what appears to be such a large drop in her grades should worry Kendra? Explain your reasoning.

____

____

____

Name ____ Date_______________ Class_____________

WORKSHEET

29 COMMUNICATING SKILLS

Making Data Mean*ingful*

The following sentences use the word average in different ways:

- He was just an ordinary, average guy.
- The average volume of the six solids was 3.2 cm^3.

1. What is different about the way *average* is used in each sentence?

____

____

2. What is similar about the way *average* is used in each sentence?

____

____

What Does It All *Mean*?

Because average can be used in different ways, scientists use the word **mean** instead. In this sense, *mean* is the same as a mathematical average. For instance, to find the mean height of seven students, you add up their individual heights and divide the sum by seven, the number of students.

Suppose the seven students above are third-graders who live in Charlotte, North Carolina. If you wanted to find the mean height of third-graders in Charlotte, you could do one of the following two things:

- You could measure the height of every single third-grader in Charlotte, and then calculate the *population mean.* This would take a long time because there are thousands of third-graders in Charlotte. The **population mean** refers to a mathematical average that has been calculated based on *all* of the available data.
- You could measure the height of several third-graders in certain areas and calculate the *sample mean.* The **sample mean** refers to a mathematical average that has been calculated based on only *some* (a sample) of the available data. The sample mean is an estimate of the population mean.

3. When do you think it is more appropriate to calculate a sample mean? Can you think of any problems with using a sample mean?

____

____

4. When do you think it is more appropriate to calculate a population mean? Can you think of any problems with using a population mean?

____

____

Mode, Median, and Range

Mode, median, and range are other important mathematical tools for interpreting data. The **mode** is the value that occurs most often in a set of data. For example, imagine you are counting the number of slices of pepperoni on certain pieces of pizza. Following are the outcomes for 10 pieces:

Number of Pepperoni Slices per Piece of Pizza

Pizza piece	A	B	C	D	E	F	G	H	I	J
Pepperoni slices	5	5	2	3	2	5	6	1	2	5

Which number occurs most often in the data set? By counting how many times each number appears, you find that 5 appears most often. Therefore 5 is the *mode*. It is possible to have more than one mode in a set of data. If two values *tie* for the most occurrences, the data set has two modes. For example, if there had been one more 2 in the data above, the modes would have been 5 and 2.

The **median** is the middle value of a set of data listed in numerical order. If a set of data contains an even number of items, it will have two middle numbers. In this case, to find the median, you average the two middle numbers.

Let's find the median for the data set listed above. First, put the data in numerical order, from least value to greatest value.

1 2 2 2 3 5 5 5 5 6

Notice there are two middle numbers in this set, 3 and 5. We must average them then to get the median; $(3 + 5) \div 2 = 4$, so 4 is the median!

The **range** is the difference between the greatest number and the smallest number in the data set. Range shows how much the data set varies. Let's find the range of the set of data above. The greatest number is 6, and the smallest number is 1. To find the difference, we subtract 1 from 6.

$$6 - 1 = 5$$

The range for the data set is 5, so the data vary over 5 values.

Your Turn

Find the statistical measures for the following sets of data:

5. 3, 5, 7, 9, 5, 6, 4, 3, 2, 22

a. Mean

____

b. Mode

____

c. Median

____

d. Range

____

6. 4, 19, 3, 19, 4

a. Mean

____

b. Mode

____

c. Median

____

d. Range

____